

分散分析における変量効果モデル － 心理学研究への適用について －

菅 千 索

(心理学教室)

Random-effect Model in Analysis of Variance － On Application to Psychological Studies －

Sensaku SUGA

2003年10月2日受理

はじめに

心理学研究で使われている実験計画 (experimental design) においては、独立変数 (independent variable) である要因 (factor) の各水準 (level) に対して、異なる被験者を割り当てる被験者間配置 (between-subject design) と、同じ被験者で繰り返し測定する被験者内配置 (within-subject design) という2つの区別がある。そして1つの実験計画のなかに、少なくとも1つ以上の被験者内配置となる要因 (以下、これを「被験者内要因」とよぶ) が存在する場合は、1人の被験者につき2つ以上の測定値が得られているため、被験者に関する個人差の問題が研究目的に含まれているかに関わりなく、分散分析 (analysis of variance) では被験者そのものが新たな1つの要因として扱われねばならないことになる (以下、これを「被験者要因」とよぶ)。たとえば、調べたいのは1つの実験要因だけであるが、これを被験者内配置で実施すれば、分散分析は2要因 (実験要因×被験

者要因) で行わなくてはならないのである。このような被験者内要因が含まれている実験計画なかでも、すべての要因が被験者内配置であるものは乱塊法 (randomized block design)、被験者間配置と被験者内配置が混在しているものは入れ子配置 (nested design) または枝分かれ配置とよばれている。また、被験者内配置をまったく含まないのが完全無作為化法 (completely randomized design) である (以下、被験者間配置となる要因を「被験者間要因」とよぶ)。

一方、この被験者間配置と被験者内配置とは別に、要因には各水準がどのように定められたかに応じて、固定効果モデル (fixed-effect model) と変量効果モデル (random-effect model) という異なる理論上の仮定が存在する。奥野・芳賀 (1969) によれば、前者は『その水準の比較を目的とし、特定の水準だけを取り上げるときのモデル』、後者は『その水準は、不特定多数の集合の中から無作為に取り出されたものと考え、したがって、特定の2水準の比較をすることは無意味で、水準

効果の分散の大きさだけに関心があるときのモデル』と説明されている。ある特定の要因をどちらのモデルと仮定して分析するかによって生じる数学的な違いは、その要因が関わる変動因（主効果と交互作用）の平均平方（mean square）の期待値に現れることになる。そして固定効果モデルではなく変量効果モデルとして扱うことでの実際上の影響は、重要と思われる順に(1)検定ができなくなる、(3)検定ができるための条件が厳しくなる、(3)事後検定である多重比較が行えない、(4)検定の方法が異なる、(5)検定の前提となる仮定が違う、(6)特に注意すべき点はない、と整理できる。

ほとんどすべての心理学研究において、研究者が研究目的に応じて取り上げる要因というのは、つねに固定効果モデルであると考えてよい。なぜならば、そこでの各水準は先行研究や独自の予備実験・調査、または少なくとも主観的な経験や予想にもとづいて定められており、決して母集団からの無作為抽出ではないからである。たとえば、記銘時間の要因が再生量に及ぼす記憶実験を行う際に、短時間条件として3分間、長時間条件として9分間を設定されたとする。なぜ3分間と9分間でなければならないのか、または、2分間と8分間ではだめな理由はあるのか、などさまざまな疑問が生じるかもしれないが、それらに対して説明や反論ができなくても、研究者がとにかく3分間と9分間の比較をしたいと考えている限りは固定効果モデルなのである。別の言い方をすれば、3分間と9分間の選択に関する適切性または妥当性というのは、実験仮説を提示するまでの論理展開のなかで示されるべきものであって、実験計画・分散分析という枠組み以前の問題だと考えるべきなのである。一方、ここでおそらくは唯一の

例外だと考えておかねばならないのは、直接かつ主要な研究目的のなかに、個人差の一般的な存在が含まれる場合であろう。そこでは特定の個人間の比較を行うだけでは不十分であり、被験者は想定される母集団のなかから無作為に抽出された標本でなければならないため、当然ながら変量効果モデルが採用されなければならないと一般には考えられている（この点についての反証は後述する）。

さて、本論で議論しようとしているのは、各被験者が各水準に対応するような被験者要因のなかでも、個人差そのものが主たる研究目的に含まれているのではなく、最初に述べたように実験計画の段階で何らかの意図から被験者内配置とされた結果、新たに被験者そのものを要因として追加せねばならない場合についてである。この場合においても、おそらくは「被験者は母集団から無作為抽出された標本であるから、被験者個人ごとを水準とする要因は変量効果モデルでなければならない」という理由から、そのような扱いをするのが心理学研究における従来からの通説になっているようにも思われる。そこで心理学研究において一般に使われていると考えられる統計法あるいはデータ分析法に関する書籍のなかで（ただし筆者が所有しているものに限る）、実験計画または分散分析について一定以上の記述があるものをレビューしてみた。その結果、両モデルについての一定以上の詳しい解説、および被験者要因に対して仮定するモデルが異なるとどうなるかについての具体的な検討が行われているのは岩原（1965）と生沢（1977）だけであった。また、両モデルに関するやや詳しい解説だけがあったのは肥田野・瀬谷・大川（1961）と森・吉田（1990）であり、具体的な分析方法に関しては、前者

は固定効果モデルだけしか解説していなかったのに対して、後者は基本的に被験者要因は変量効果モデルとする立場をとっていた。さらに、両モデルの簡単な説明だけにとどまっていたのは、繁桙・柳井・森（1999）、倉智・山上（1991）、海保（1985）、後藤・大野木・中澤（2000）、南風原（2002）であった。逆に、両モデルにはほとんど触れず、具体的な分析方法の紹介では被験者要因を一貫して変量効果モデルとして扱っていたのは篠原（1984a, 1984b, 1986）、田中・山際（1989）、田中（1996）であり、これらの問題にはほとんど触れていなかったのは山内（1998）、渡辺（2002）、岩原（1967）、遠藤（2002）であった。

用語の再検討

実験計画と分散分析 実験計画は1920年代に R. A. Fisherによって農業試験の分野で基礎が築かれたのち、とくに第2次世界大戦後は理工農学系だけでなく医学・薬学、生物学、心理学などの分野でも広く使われるようになった。その目的は、正確かつ有意義な情報を効率的に収集するのに最適な実験方法を考案することにある。そこでは実験仮説の検証に必要な要因および水準の選択方法と、データに混入する有害な誤差の制御という2点がとりわけ重要である。一方、分散分析は数理統計学の応用分野だと位置づけられ、実験によって得られたデータを解析する数学的手法を扱っている。このように書くと両者は互いに異なる学問領域のように思われるかも知れないが、実際にはコインの表裏のような不可分な関係にある。なぜならば、分散分析によって数学的に解析可能なデータが収集されなければ実験を行う意味はなく、逆にいえば、不

適切な実験計画から得られたデータは解析不可能な場合も少なくない。したがって、両者の間には、実験計画は文字通り実験を計画する段階での議論であり、分散分析は実験実施後の使用される手法であるといったニュアンスの違いがある程度だと考えておけばよいであろう。そこで本論では両者を特に区別せず、文脈に応じて使い分けることにする。

従属変数と独立変数 分散分析における従属変数（dependent variable）とは、実際に測定が行われる変数を指すもので、心理学研究においては、たとえば学力テストの得点や刺激への反応時間、記憶の正再認率など実験の結果そのものを表す数値であり、測定値や特性値などと呼ばれることもある。通常の実験計画における従属変数は、間隔尺度以上の尺度水準を満たす量的変数でなければならず、かつ一度の分散分析に含まれる従属変数は常に1つだけである。この従属変数の変動が、ほかのどのような変数とどのように関係しているかに関して、先行研究の結果や理論的な考察などにより選ばれたものを独立変数という。従属変数が実験結果だとすれば、独立変数は実験条件だと考えてもよく、満たされるべき尺度水準は名義尺度である。一般に上位の尺度水準は下位の水準の情報も包含しているため、独立変数は離散変数として扱うことができればどのような変数でもよく、しかも1度の分散分析に複数の独立変数が含まれてもよい。ちなみに独立変数がすべて間隔尺度以上である場合は、分散分析ではなく回帰分析を利用の方が数学的には望ましい。また、理工農学系では独立変数のことを原因系、従属変数のことを結果系とよぶことがあるが、心理学研究においては因果関係を判断するこ

とが容易でないため、あまり使わない方がよいと考えられる。

要因と水準 実験計画では独立変数そのもののことを要因、そして独立変数が実際にとる値のことを水準という。たとえば性差要因における男子水準と女子水準、または学習時間要因における長時間水準と中時間水準と短時間水準などであり、実験が行われる条件の違いが水準だと考えてよい。そして実験計画において1つの要因が成立するためには、その要因には2つ以上の水準が存在し、かつ、それぞれの水準には2個以上のデータが含まれていなければならない。なお、このように1つの水準内で複数の測定を行うことを実験計画では繰り返し (repetition) という。この要因が成立するための条件を逆からみれば、複数の水準になりうるものがあり、そのそれぞれに2個以上のデータがあれば要因として取り上げねばならないことになる。すでに述べたように、少なくとも1つの被験者内要因が実験計画のなかに存在するならば、被験者を水準とする要因を新たに追加せねばならないのである。

1 要因計画と多要因計画 分散分析における従属変数は常に1つであり、独立変数である要因は単独でも複数でもよいことはすでに述べたが、この点についてもう少し掘り下げて検討してみたい。まず実験で比較したいのは1要因だけであり、これを被験者間配置で実施するときは、各水準に2名以上の被験者を割り当てる完全無作為化法の1要因計画となる。一方、これを被験者内配置で実施する乱塊法であれば被験者そのものを要因に追加した2要因計画と考えねばならない。一般に2

要因以上となる多要因計画において、要因の各水準が交差した部分を処理 (treatment) または細胞 (cell) という。ただし1要因計画の場合は水準と処理は同じものを指していることになり、少なくとも1つの水準には2つ以上のデータがなければならぬため、水準(処理)内のデータ数(繰り返しの数)が複数で等しい場合と異なる場合の2通りが考えられる。一方、多要因計画において各処理ごとに複数の測定を行うことを反復 (replication) というが、反復が行われているものと、そうでないものに分類することができ、さらに前者は処理内のデータ数(反復の数)が等しい場合と異なる場合に分けられる。

主効果と交互作用 分散分析の前半部分というのは、従属変数の変動(分散)を独立変数と対応づけて分解していくプロセスといえるが、その際にそれぞれの独立変数ごとの変動を主効果 (main effect)、主効果による変動を減じた上で複数の独立変数の組み合わせに起因する変動を交互作用 (interaction)、そして両者を総称して要因効果 (factorial effect) とよぶ。さらに交互作用には主効果を差し引いた2要因交互作用または1次交互作用 (primary interaction) のほか、主効果と2要因交互作用を差し引いた上での3要因交互作用または2次交互作用 (secondary interaction)、さらに同様の考え方で導かれる高次の交互作用が考えられる。これらの要因効果をすべて分解した上で残る変動を誤差 (error) または残差 (residual) いい、通常は要因効果を評価する場合の基準として使われている。ただし、多要因計画で実際に誤差が求められるのは細胞内に複数のデータが存在する場合に限られるため、細胞内のデータ数が1つの場合には

最高次の交互作用を誤差とみなして代用することになる。

被験者間配置と被験者内配置 心理学研究において実験計画を立てる際、被験者のもつ何らかの特性の違いを水準とする要因を考える場合は、同じ被験者をすべての水準に割り当てる方法（被験者内配置）は不可能であるから、各水準に異なる被験者を割り当てる被験者間配置を用いなければならない。ここでいう特性の違いとは、たとえば性別の差、居住地域の差、所属組織・学校の差、性格特性や諸能力の差などである。さらに、ある水準で1度測定されてしまうと、その影響が学習や疲労・飽和などで別の水準において同じ被験者に残る場合も被験者間配置によらなければならないことが多い。極端な例ではあるが、偶発学習課題の場合は1度やってしまうと「ネタ」がばれてしまうし、瞬間視で1度でも知覚した刺激はそれ以後は見えやすくなる。以上は被験者間配置でなければならないか、またはその方が明らかに望ましいと考えられるケースであったが、逆に被験者内配置の方が合理的な、または有効な場面も考えられる。たとえば、何らかの指導や訓練などの経験をした前後で被験者にどのような変化がみられたかをしらべるときは、被験者間配置でやれば必要な収集したいデータ数の2倍の被験者を用意せねばならないので、同じ被験者で事前と事後に2度測定する被験者内配置の方がはるかに合理的である。また、実験の目的である要因効果に比べて被験者の個人差の方が相対的にある程度以上大きいと考えられる場合は、被験者間配置で被験者を無作為に割り当てるよりも、同じ被験者の中での変化を調べる方が有効だと考えられることも少なくな

い。したがって、すでに述べた学習や疲労・飽和などに十分注意し、必要ならば実施方法などでそれらへの対応をした上であれば、被験者内配置を採用するメリットが大きいケースも十分に考えられるであろう。なお、本論ではある要因が個人内配置とされた結果として現れた各被験者を水準にとる要因を特に被験者要因と呼ぶことにしているが、当然これは被験者間配置の一形態に他ならない。

固定効果モデルと変量効果モデル 両モデルの違いに関する一般的な記述については奥野・芳賀（1969）を引用したが、さらに被験者要因をどちらのモデルを扱うべきかに関しては本論の中心的な検討課題であるため、加えて生沢（1977, pp.64-65）の記述を詳しく引用しておきたい。そこで例として示されているのは「大学生3名に、カード（いわゆるトランプ）52枚を、赤黒に分類させた。分類に要した時間を、各人4回ずつ測定し、表4-4（本論では省略）の結果を得た（大阪市大実験レポート）。」という状況下であって、「被験者として用いたその3名だけについて検討したい。」ときは固定効果モデル、「その3名は、母集団からの無作為抽出標本と考える。データは3名分だが、検定は母集団全体としてバラツキがあるかどうかを判断したい。」ときは変量効果モデルであると区別している。さらに生沢は被験者に限らず両モデルの一般的な違いについても触れているが、その最後に「その中間で、 p 個の条件は、多数の中からたまたま選ばれたものだが、変量効果モデルとも思われぬ場合は、固定効果モデルとして扱わざるを得ない。ただし、検定は p 個かぎりに関するものである」と付け加えている。なお、表記法の問題として、固定効果モデルは単に固定

モデル (fixed model) や定数 (効果) モデルや母数モデル、また変数効果モデルは単に変数モデル (random model) や無作為モデルとよばれることもある。また「モデル」の代わりに「模型」という訳語が使われることも多い。さらに、そのモデルが仮定されている要因という意味で固定効果要因 (因子または変数ともいう) や変数効果要因 (同左) と表記することもあるが、これらはすべて同義であると考えても問題はない。

平均平方の期待値計算

本論では異なるタイプの分散分析を幾つか論じるが、それらすべてについて平均平方の期待値の計算式を示すことは現実的に不可能であるし、また、あまり有益であるとも思われない。しかしながら、固定効果モデルと変数効果モデルの違いを検討する上では決して避けて通れない問題でもあるため、ここではもっとも簡単な固定効果モデルによる 1 要因完全無作為化法についての計算式 (証明) だけを明示しておくことにしたい。

まず、ここで扱うタイプの分散分析におけるデータの構造モデルは、

- x_{ij} : A_i 水準 ($i=1, 2, \dots, a$) の j 番目の測定値
 μ_i : A_i 水準の真の値
 e_{ij} : 測定値 x_{ij} に含まれる測定誤差; N.I.D ($0, \sigma^2$)

とするとき、

$$x_{ij} = \mu_i + e_{ij}$$

と表される。そこで

$$\mu = \bar{\mu} = \frac{1}{a} \sum_{i=1}^a \mu_i \quad (\text{全平均})$$

$$\alpha_i = \mu_i - \mu \quad (\text{要因} A \text{ の主効果 ; } \sum_{i=1}^a \alpha_i = 0)$$

とおけば、

$$\mu_i = \mu + (\mu_i - \mu) = \mu + \alpha_i$$

であるため、構造モデルは最終的に

$$x_{ij} = \mu + \alpha_i + e_{ij}$$

と定式化される。ちなみに、固定効果モデルの場合には上式のように要因 A の主効果の総和をゼロと置くことができるが (「一般平均のほうにしわ寄せすることにより水準効果の和は 0 と考えることができる (鷺尾, 1974, p.43)。」)、変数効果モデルならば母集団の一部にすぎない標本であるため総和 0 とならない点、両モデルの違いを特徴づけるポイントとなるのである。

つぎに準備のために水準および全体の平均を構造モデルで表しておけば、

$$\bar{x}_r = \frac{1}{r} \sum_{j=1}^r x_{rj} = \frac{1}{r} \sum_{j=1}^r (\mu + \alpha_r + e_{rj}) = \mu + \alpha_r + \bar{e}_r$$

$$\bar{x}_{..} = \frac{1}{a} \sum_{i=1}^a \bar{x}_i = \frac{1}{a} \sum_{i=1}^a (\mu + \alpha_i + \bar{e}_i) = \mu + \bar{e}_{..}$$

となる。さて、ここでの分散分析表を先に示しておく (表 1)。

表 1 1 要因完全無作為化法の分散分析表 (公式)

変動因	平方和	自由度	平均平方	F 値
要因 A	$SS_A = \sum_{i=1}^a \sum_{j=1}^r (\bar{x}_i - \bar{x}_{..})^2$	$df_A = a - 1$	$MS_A = SS_A / df_A$	$F_A = MS_A / MS_R$
残差 R	$SS_R = \sum_{i=1}^a \sum_{j=1}^r (x_{ij} - \bar{x}_i)^2$	$df_R = a(r - 1)$	$MS_R = SS_R / df_R$	
全体 T	$SS_T = \sum_{i=1}^a \sum_{j=1}^r (x_{ij} - \bar{x}_{..})^2$	$df_T = ar - 1$		

まず要因 A の平均平方の期待値は、

$$\begin{aligned}
 E\{MS_A\} &= E\left\{\frac{1}{a-1}\sum_{i=1}^a\sum_{j=1}^r(\bar{x}_r - \bar{x}_{..})^2\right\} \\
 &= \frac{r}{a-1}E\left\{\sum_{i=1}^a(\bar{x}_r - \bar{x}_{..})^2\right\} \\
 &= \frac{r}{a-1}E\left\{\sum_{i=1}^a[(\mu + \alpha_i + \bar{e}_r) - (\mu + \bar{e}_{..})]^2\right\} \\
 &= \frac{r}{a-1}E\left\{\sum_{i=1}^a(\alpha_i + \bar{e}_r - \bar{e}_{..})^2\right\} \\
 &= \frac{r}{a-1}E\left\{\sum_{i=1}^a\alpha_i^2 + 2\sum_{i=1}^a\alpha_i(\bar{e}_r - \bar{e}_{..}) + \sum_{i=1}^a(\bar{e}_r - \bar{e}_{..})^2\right\} \\
 &= \frac{r}{a-1}\left[E\left\{\sum_{i=1}^a\alpha_i^2\right\} + 2\sum_{i=1}^a\alpha_i E\{\bar{e}_r - \bar{e}_{..}\} + E\{(\bar{e}_r - \bar{e}_{..})^2\}\right]
 \end{aligned}$$

とまで計算して、これに

$$\begin{aligned}
 E\left\{\sum_{i=1}^a\alpha_i^2\right\} &= \sum_{i=1}^a\alpha_i^2, \\
 E\{\bar{e}_r - \bar{e}_{..}\} &= 0, \quad E\{(\bar{e}_r - \bar{e}_{..})^2\} = (a-1)Var\{\bar{e}_r\} = (a-1)\frac{\sigma^2}{r} \quad (i=1, 2, \dots, a)
 \end{aligned}$$

を代入すれば、結局は

$$E\{MS_A\} = \frac{r}{a-1}(a-1)\frac{\sigma^2}{r} + \frac{r}{a-1}\sum_{i=1}^a\alpha_i^2 = \sigma^2 + \frac{r}{a-1}\sum_{i=1}^a\alpha_i^2$$

と求められる。

一方、同様にして残差項の平均平方の期待値を求めてみると、

$$\begin{aligned}
 E\{MS_R\} &= E\left\{\frac{1}{a(r-1)}\sum_{i=1}^a\sum_{j=1}^r(x_{ij} - \bar{x}_r)^2\right\} \\
 &= \frac{1}{a(r-1)}E\left\{\sum_{i=1}^a\sum_{j=1}^r[(\mu + \alpha_i + e_{ij}) - (\mu + \alpha_i + \bar{e}_r)]^2\right\} \\
 &= \frac{1}{a(r-1)}E\left\{\sum_{i=1}^a\sum_{j=1}^r(e_{ij} - \bar{e}_r)^2\right\} \\
 &= \frac{1}{a(r-1)}\sum_{i=1}^aE\left\{\sum_{j=1}^r(e_{ij} - \bar{e}_r)^2\right\}
 \end{aligned}$$

と展開し、そこへ

$$E\left\{\sum_{j=1}^r(e_{ij} - \bar{e}_r)^2\right\} = (r-1)Var\{e_{ij}\} = (r-1)\sigma^2 \quad (i=1, 2, \dots, a)$$

を代入すれば、

$$E\{MS_R\} = \frac{1}{a(r-1)}a(r-1)\sigma^2 = \sigma^2$$

が得られるのである。

この節では要因の 1 つが被験者要因となる場合であって、その要因を固定効果モデルとするか変量効果モデルとするかで平均平方の期待値がどのように変わるか、さらに、その

分散分析ごとの理論的考察

結果として検定方法や前提条件がどう違って
いるかを論じることにより、両モデルの特徴
について考察する。なお、期待値の表記は生
沢 (1977) に準じているが、そこで扱われて
いないものについては岩原 (1965) を参考に
している。

1 要因で細胞の大きさが 2 以上の分散分析

これは一般に 1 要因完全無作為化法とよばれ
ているもので、ここでの 1 要因は被験者間配
置でなければならない。この場合の平均平方
の期待値は表 2 となる。

表 2 平均平方の期待値 (1)

変動因	期待値
要因 A	固定: $\sigma^2 + \frac{r}{a-1} \sum_{i=1}^a \alpha_i^2$ 変量: $\sigma^2 + r\sigma_A^2$
残差 R	σ^2

注: 「固定」と「変量」の欄は要因 A をどちらと仮定す
るかの意味である。

すなわち、固定効果モデルでは主効果の不
偏分散であった部分が、変量効果モデルでは
母分散に置き換えられているだけで、計算や
検定の方法には両モデル間で何も違いはない。
ただし生沢 (1977) によれば、変量効果モデ
ルにおいては、各処理のデータ数は等しくな
なければならないという制約が生じるという。

このタイプの分散分析で要因 A を変量効果
モデルとして扱わねばならないのは、被験者
間要因のなかでも各水準に個々の被験者が対
応している被験者要因だけだと考えてよい。
つまり個人差の一般的な存在を検証している
ことになるが、ここでの帰無仮説は被験者の
母集団分布にバラツキがない、すなわち母集
団に含まれるすべての個体は全部同じである
とされている。したがって、この帰無仮説が
棄却されたとしても、多数の個体のなかに 1
人だけが特異値となるような状況も考慮して
おかねばならないため、心理学研究において
はほとんど意味がない実験仮説だといわざる

を得ない。また、各水準内で同じ被験者を反
復測定せねばならないが、このことは水準内
での無作為化を困難にしているし、測定値(測
定誤差)は互いに独立であるという前提条件
も疑わしくなっている。これらの議論を総括
すれば、このタイプにおいて変量効果モデル
を採用することの長所あるいは利点は一切な
く、逆に分析可能な条件が厳しくなったり、
実験仮説が非現実的となるだけなのである。

2 要因で細胞の大きさが 1 の分散分析

ここで 2 要因とも実験仮説である場合 (条件×条
件) については、変量効果モデルを考慮する
必要はない。一方、実験仮説は 1 要因である
が、これを被験者内要因で実施した場合 (条件
×被験者) は、このタイプの 2 要因分散分析で
解析されることになり、乱塊法または層別無
作為化法と呼ばれている。その際、被験者要
因については変量効果モデルが仮定されるの
が心理学研究では通例であるが、要因 B を被

験者要因とみなして、これに両モデルを仮定したときの平均平方の期待値を表3に示す。

表3 平均平方の期待値(2)

変動因	期待値
要因A	固定: $\sigma^2 + \frac{b}{a-1} \sum_{i=1}^a \alpha_i^2$ 変量: $\sigma^2 + \frac{a}{a-1} \sigma_{AB}^2 + \frac{b}{a-1} \sum_{i=1}^a \alpha_i^2$
要因B	固定: $\sigma^2 + \frac{a}{b-1} \sum_{j=1}^b \beta_j^2$ 変量: $\sigma^2 + a\sigma_B^2$
残差R	固定: σ^2 変量: $\sigma^2 + \frac{a}{a-1} \sigma_{AB}^2$

注:「固定」と「変量」の欄は要因Bをどちらと仮定するかの意味である。

まず、ここでの残差Rは、細胞の大きさが1であるから測定誤差が推定できないために、その誤差と要因A×Bの交互作用が交絡したものであることに注意せねばならない。そのために2要因がともに固定効果モデルの場合には、分析の前提条件として「交互作用は有意でない」を課した上で、残差Rを誤差項(F比の分母)にして2つの主効果を検定することになる。また、要因Bを変量効果モデルと仮定する場合であっても、やはり「交互作用は有意でない」と仮定するのであれば要因Aおよび残差Rの第2項は消去されるため、まったく同じ方法で2要因の主効果が検定できることになり、要因Bをどちらとみなすかの実質的な違いはない。

一方、要因B(被験者要因)を変量効果モデルと仮定した場合に限っては、「交互作用が有意でない」という前提条件を設けなくても、要因Aの主効果だけは検定可能となる(要因Bの主効果の検定は不可能)。ただし、その場合であっても「単純交互作用の分散はみな等

しいこと(岩原、1965, p.253)」という前提条件は残ってしまうが、そのような仮定が妥当であるかについては、非常に予測が困難だと考えられる。さらにいえば、交互作用が有意であるということは、要因Aの水準の違いがもつ効果が被験者ごとに異なる訳であるから、そもそも被験者全体を込みにした要因Aの主効果を検定する意義はすでに薄れているといえる。

したがって、『被験者要因(要因B)を変量効果モデルにすれば交互作用があっても要因Aの主効果が検定できる』という表現はもっともらしいが、詳しく検討してみると現実にはかなり無理があるように思われる。このタイプの分散分析は前提条件として交互作用が有意でないとの仮定が妥当なときだけの用いるべきであり、その場合は被験者要因が固定効果モデルか変量効果モデルかという議論をする必要はない。また、固定効果モデルを仮定することで、局所管理という立場から被験者の個人差についても言及することも可能に

なる。それに対して、「交互作用が有意でない」との仮定に無理がある場合には、データ解析の段階で変量効果モデルを仮定して対応するのではなく、交互作用が検出できるような実験計画を立てるのが本筋であろう。

2 要因で細胞の大きさが 2 以上の分散分析

2 要因ともが被験者間配置の実験要因であれば、一般に 2 要因完全無作為化法とよばれて

いるもので、その場合は当然ながら 2 要因とも固定効果モデルだと考えても問題はないであろう。したがって、検討しておかなければならないのは『1 要因で細胞の大きさが 2 以上の分散分析』ですでに述べたように、要因の 1 つが被験者要因のときについてである。そこで要因 B を被験者要因とみなして、これに両モデルを仮定したときの平均平方の期待値を表 4 に示す。

表 4 平均平方の期待値 (3)

変動因	期待値	
要因 A	固定: $\sigma^2 + \frac{br}{a-1} \sum_{i=1}^a \alpha_i^2$	変量: $\sigma^2 + \frac{ar}{a-1} \sigma_{AB}^2 + \frac{br}{a-1} \sum_{i=1}^a \alpha_i^2$
要因 B	固定: $\sigma^2 + \frac{ar}{b-1} \sum_{j=1}^b \beta_j^2$	変量: $\sigma^2 + ar \sigma_B^2$
交互作用 AxB	固定: $\sigma^2 + \frac{r}{(a-1)(b-1)} \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2$	変量: $\sigma^2 + \frac{ar}{a-1} \sigma_{AB}^2$
残差 R	σ^2	

注: 「固定」と「変量」の欄は要因 B をどちらとみなすかの意味である。

ここで要因 B を固定効果モデルとした場合には要因 A と B の主効果、および交互作用 AxB の検定はすべて残差 R が誤差項 (F 比の分母) となる。一方、変量効果モデルであれば要因 A の主効果では交互作用 AxB が、また要因 B の主効果および交互作用 AxB では残差 R がそれぞれ誤差項となる。すなわち、要因 B をどちらのモデルと仮定するかで要因 A の主効果を検定する方法が異なることになり、この点に注目して変量効果モデルの存在を強調する見解が一般的に行われているように思われる。しかしながら、この問題に対する奥

野・芳賀 (1969, p.116; p.237) の主張は是非とも注目しておきたい。

彼らは『しかし、もし読者が「交互作用」の意義を十分理解されるなら、上述の 2 つの検定法の相違は、実質上あまり意義をもたないことがわかるであろう (p.236)。』と述べた上で、次のように主張している。まず、交互作用 AxB が有意でない場合には、交互作用 AxB と残差 R の平均平方はあまり変わらないため、要因 A の主効果の誤差項をどちらにしても大差はないという。もちろん、実際の数値計算の段階で有意水準 5% にクリティカル

な場合には、どちらを分母にとるかで検定結果がことなることもあるが、ここで述べられているのは実験計画と分散分析に対する基本姿勢だと理解すべきであろう。一方、交互作用AxBが有意であれば、そもそも要因Aの主効果を検定することの意義は薄れ、事後検定として単純主効果の検定を行う方が望ましいとされているため(たとえば森・吉田(1990)など)、誤差項を交互作用AxBにするか残差Rにするか議論はあまり意味がないことになる。ただし、交互作用AxBが有意であっても、その分散(平均平方)と比べて要因Aの分散がはるかに大きいときは要因Aの主効果にも注目すべきであるが、この時点で変量効果モデルと同様に要因Aと交互作用の比較も行っていることになる。したがって、2要因とも固定効果モデルだと仮定しても、分析結果の解釈が正しく行われるならば、まったく問題がない考えられるという。さらには、要因Bで変量効果モデルを仮定した場合は、交互作用

は正規分布に従い、かつ単純交互作用の分散はすべて等しいという前提条件がつくことになるが、これらは非現実的な条件だと認識しておかねばならない。なお、このタイプの分散分析において、まず最初に交互作用AxBの有意性の検定を行い、それが有意であれば要因Aの主効果の誤差項は交互作用AxB、有意でなければ誤差項を残差Rとするという方法も考えられている。

3 要因で細胞の大きさが1の分散分析 これまで3つの分散分析の検討を通して基本的な問題点は指摘し終えたと思われるので、以下では同様のことが他の分散分析でも見られることを簡潔に示しておきたい。さて、この分散分析において要因Cが被験者要因であれば、要因AおよびBはともに被験者内配置ということになり、一般に2要因乱塊法とよばれているものに相当する(表5)。

表5 平均平方の期待値(4)

変動因	期待値	
要因A	固定: $\sigma^2 + \frac{bc}{a-1} \sum_{i=1}^a \alpha_i^2$	変量: $\sigma^2 + \frac{ab}{a-1} \sigma_{AC}^2 + \frac{bc}{a-1} \sum_{i=1}^a \alpha_i^2$
要因B	固定: $\sigma^2 + \frac{ac}{b-1} \sum_{j=1}^b \beta_j^2$	変量: $\sigma^2 + \frac{ab}{b-1} \sigma_{BC}^2 + \frac{ac}{b-1} \sum_{j=1}^b \beta_j^2$
要因C	固定: $\sigma^2 + \frac{ab}{c-1} \sum_{k=1}^c \gamma_k^2$	変量: $\sigma^2 + ab\sigma_C^2$
交互作用 AxB	固定: $\sigma^2 + \frac{c}{(a-1)(b-1)} \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2$	変量: $\sigma^2 + \frac{ab}{(a-1)(b-1)} \sigma_{ABC}^2 + \frac{c}{(a-1)(b-1)} \sum_{i=1}^a \sum_{j=1}^b (\alpha\beta)_{ij}^2$

$$\text{交互作用 AxC 固定: } \sigma^2 + \frac{b}{(a-1)(c-1)} \sum_{i=1}^a \sum_{k=1}^c (\alpha\gamma)_{ik}^2 \quad \text{変量: } \sigma^2 + \frac{ab}{a-1} \sigma_{AC}^2$$

$$\text{交互作用 BxC 固定: } \sigma^2 + \frac{a}{(b-1)(c-1)} \sum_{j=1}^b \sum_{k=1}^c (\beta\gamma)_{jk}^2 \quad \text{変量: } \sigma^2 + \frac{ab}{b-1} \sigma_{BC}^2$$

$$\text{残差 R} \quad \text{固定: } \sigma^2 \quad \text{変量: } \sigma^2 + \frac{ab}{(a-1)(b-1)} \sigma_{ABC}^2$$

注:「固定」と「変量」の欄は要因Cをどちらとみなすかの意味である。

まず3要因とも固定効果モデルだと仮定すれば、2次の交互作用AxBxCは有意でないという前提条件のもとで、残差Rを誤差項として3つの主効果と3つの一次の交互作用の検定を行うことになる。一方、要因Cを変量効果モデルと仮定した場合は、要因Aの主効果は交互作用AxC、要因Bの主効果は交互作用BxC、交互作用AxBは残差R(交互作用AxBxC)をそれぞれ誤差項として検定を行うことになる。その上で2次の交互作用AxBxCが有意でないと仮定できるのであれば、要因Cと交互作用AxCと交互作用BxCは残差Rを誤差項として検定できるが、そうでなければ検定できないことになる。

このように表現すると要因Cにどちらのモデルを仮定するかによって、分散分析に大きな違いが現れるという印象を与えるかもしれないが、すでに述べたように交互作用の意味をよく理解しているならば、実質的な違いがないことが理解できるであろう。さらに、ここで重ねて指摘しておきたいのは、変量効果モデルを仮定したときに誤差項となる交互作用AxCとBxC、残差R(交互作用AxBxC)はすべて正規分布に従い、さらに単純交互作用の分散は等質でなければならないという仮定は、極めて現実的ではないという点である。

1 要因で枝分かれ配置となる分散分析(欠損でない細胞の大きさ2以上) 枝分かれ配置というのは、すべての処理についてはデータが収集されていないタイプの実験計画のことであり、不完備型計画ともよばれている。たとえば、ここでの例として要因Aが2水準だとして6名の被験者を3名ずつ割り当てる場合を考えてみる。もしも、各被験者から1つのデータが得られるのであれば普通の1要因完全無作為化法である。一方、繰り返し測定して各被験者から複数のデータを得たとすれば、被験者も要因として考える必要があるため、見かけ上は2×6の2要因計画となる(事前に被験者ごとの複数のデータの総和または平均を求めてから分散分析を行うのであれば、1要因完全無作為化法に帰着するが、その場合は被験者の主効果については検定できない)。しかしながら、水準1に割り当てられた3名については水準2でのデータはなく、残りの3名についてもその逆のことがいえる。すなわち、1要因乱塊法(条件×被験者)ならばすべて揃っているはずのデータが、ここではそもそも得られないような実験計画になっているのである。このような枝分かれ配置の場合には、いくつかの変動因は分離不能

な交絡した状態でしか求められないことになり、ここでは2つの主効果と交互作用がそのようになっている。すなわち、1人目の被験者の平均値が大きかったとしても、それが要因Aの水準1の影響によるものか、それとも被験者個人の特徴によるものか、または両者

の交互作用によるものか区別しようがないのである。このタイプの分散分析には以上のような特徴はあるが、ここでの議論については、すでに述べてきたことがほぼ当てはまるので、平均平方の期待値を以下に示す（表6）。

表6 平均平方の期待値（5）

変動因		期待値
<hr/>		
要因A	固定： $\sigma^2 + \frac{sr}{a-1} \sum_{i=1}^a \alpha_i^2$	変量： $\sigma^2 + r\sigma_s^2 + \frac{sr}{a-1} \sum_{i=1}^a \alpha_i^2$
被験者	固定： $\sigma^2 + \frac{r}{a(s-1)} \sum_{i=1}^a \sum_{j=1}^s (\alpha\delta)_{ij}^2$	変量： $\sigma^2 + r\sigma_s^2$
残差R	σ^2	

注：「固定」と「変量」の欄は被験者をどちらとみなすかの意味である。

2要因とも固定効果モデルだとすれば、残差Rを誤差項にして2つの主効果を検定することになるが、被験者を変量効果モデルだとすれば要因Aの主効果は被験者を、また被験者の主効果は残差Rをそれぞれ誤差項として検定を行わねばならない。まず、これまでの議論に従えば、被験者の主効果の検定方法は実質的に同じであるから、それが有意でないときには要因Aの主効果の検定で誤差項をどちらにするかについての議論は、あまり大きな意味をもたないといえる。一方、被験者の主効果が有意であるとすれば、処理（細胞）内での個人内変動に比べて水準内での個人間変動が十分に大きいことになる。その際、処

理内での個人内変動が正規分布に従うという仮定は受け入れやすいが、それより有意に大きい水準内での個人間変動もまた正規分布に従うという仮定にはかなりの無理があり、それを誤差項として要因Aの主効果の検定を行うことは無謀だといえるであろう。

2要因で枝分かれ配置となる分散分析（欠損でない細胞の大きさが1の場合） これは被験者間配置が1要因、被験者内配置も1要因である実験計画の場合にみられる分散分析であり、その平均平方の期待値は以下の通りである（表7）。

表7 平均平方の期待値 (6)

変動因		期待値
要因 A	固定: $\sigma^2 + \frac{sc}{a-1} \sum_{i=1}^a \alpha_i^2$	変量: $\sigma^2 + r\sigma_s^2 + \frac{br}{a-1} \sum_{i=1}^a \alpha_i^2$
被験者	固定: $\sigma^2 + \frac{r}{a(s-1)} \sum_{i=1}^a \sum_{j=1}^s (\alpha\delta)_{ij}^2$	変量: $\sigma^2 + r\sigma_s^2$
要因 C	$\sigma^2 + \frac{as}{c-1} \sum_{k=1}^c \gamma_k^2$	
交互作用 AxC	$\sigma^2 + \frac{s}{(a-1)(c-1)} \sum_{i=1}^a \sum_{k=1}^c (\alpha\gamma)_{ik}^2$	
残差 R	σ^2	

注: 「固定」と「変量」の欄は被験者をどちらとみなすかの意味である。

被験者間配置である要因 A × 被験者との交互作用が交絡しているのは前項と同様である。そして被験者を固定効果モデルとすれば、残差 R を誤差項 (F 比の分母) として要因 A、被験者、要因 C の 3 つの主効果と要因 A × 要因 C の交互作用が検定できる。ちなみに残差 R は被験者 × 要因 C の交互作用と考えられるもので、さらには要因 A × 被験者 × 要因 C の 2 次交互作用はないものと仮定されている。ここで被験者を変量効果モデルと仮定すれば要因 A を検定する誤差項が残差 R から被験者に変わることになるが、それに対する検討もすでに前項で検討した通りなのでここでは繰り返さない。

な観点からの検討を通して、被験者要因は同モデルでなければならないという考え方の背景にある論理も明らかになってきたが、それらは以下の 4 点に集約されるであろう。

- (1) 被験者は無作為抽出されているはずであるから変量効果モデルが適切である。
- (2) 被験者の個人差が研究目的ならば変量効果モデルでなければならない。
- (3) 検定を行うときの誤差項が明らかに違うので、通例に従って変量効果モデルとすべきである。
- (4) 交互作用が検出できないとき、変量効果モデルならば交互作用が有意であってもよい。

まとめと結論

ここまでの変量効果モデルに対する批判的

これらについて順に考察していくことにするが、まず (1) については被験者の抽出方法が無作為的であるからといって、それが直

接的に変量効果モデルにつながる訳ではないことを確認しておきたい。すなわち、どちらのモデルを仮定するのかは、議論しようとしているのが水準が無作為抽出されたとされる母集団における分散の大きさなのか(母分散)、比較しようとして何らかの基準をもとに選ばれた水準のなかでの分散の大きさなのか(標本分散)に依存しているのである。したがって、個人内配置などにより同じ被験者から複数のデータが収集された結果として分散分析に組み入れられた被験者要因は、被験者の選択方法とは関係なく局所管理という観点も含め固定効果モデルと考えてもまったく問題はない。つまり、実際に実験に参加してくれた被験者のなかにどの程度の変動があったかというだけのことである。さらに付言しておけば、ほとんどの心理学研究における無作為抽出とは「作為的な抽出は決してやっていません」という程度に過ぎず、現実には「授業の出席者」や「無理を頼める友人・知人」、「アルバイトとしての応募者」などである。もしも分散分析において被験者要因は変量効果モデルでなければというのであれば、まずは標本抽出を厳密に行うのが先ではないであろうか。

先の(1)では個人差そのものが直接の研究対象でなく、実験計画上の必然から被験者を要因として扱わねばならない場合について述べた。それに対して、母集団における個人差を研究するときは変量効果モデルでなければならないという(2)の見解は、形式的には正しいと認めざるを得ないが、以下のような理由により実質的にはあまり意味がないと考えられる。まず変量効果モデルでは、水準は母集団から無作為に選択されているため、特定の2つの水準間の差を比較することは無

意味であり、母分散の大きさに注目せねばならない。従って、主効果の検定における帰無仮説は、固定効果モデルでは「水準の平均はすべて等しい(全体の平均と水準の平均の差はすべて0である)」となるが(これは明らかに妥当であろう)、変量効果モデルでは「母集団におけるデータの散らばり(母分散)は0である」とせねばならない。すなわち後者では、「母集団に属するすべての要素(ここでは個体)は従属変数に関してまったく同一である」というのが帰無仮説であり、その対立仮説は「少なくとも1人は従属変数に関して異なる」となる。そのため帰無仮説が棄却されて対立仮説が採択されたとしても、それは大きな母集団になかの少なくとも1人だけは「違う」という以上のことを意味していないのであり、この検定結果だけをもって「母集団には有意な個人差がある」との主張は、学問的にも現実的にも意味を持たないのである。換言すれば、母集団における個人差を議論するときには、分散分析で被験者要因を変量効果モデルとするだけではまったく不十分であり、たとえば統制条件下における個人差と実験条件下における個人差を比較するといった別途の工夫が必要である。

つぎに(3)と(4)についてであるが、2要因以上の分散分析においては、仮定するモデルによって主効果を検定するための誤差項が異なり、そのためデータが有意水準前後であれば検定結果が異なることも現実には想定されるであろう。そうした点に注意を払うことを否定するものではないが、それを強調しすぎるあまり「保守的」な通例に従って変量効果モデルを選択するというのは本筋ではない。もっとも大きな問題は、変量効果モデルをとった場合、誤差項とならねばならない

交互作用が正規分布するという前提条件が必要となってくることである。そのために交互作用は有意でないという前提条件を追加するならば固定効果モデルと同等の分析に帰着されるし、もしも有意な交互作用があってもよいというのであれば、その交互作用が正規分布に従うとの仮定、さらに、そこでの単純交互作用の等分散性の仮定は極めて困難なのである。

最後にこれまでの議論を総括すれば、少なくとも心理学研究において変量効果モデルを考える必要はなく、さらには新たな問題を引き起こす可能性が残ることが明らかになったと思われるが、もう一度だけ簡潔に整理してみると以下の3点に要約されるであろう。

① 被験者が無作為抽出されているという理由だけで、被験者要因を変量効果モデルとせねばならない必然はまったくない。モデルの選択は実験で何を実証したいのかにのみ依存している。

② 変量効果モデルの数理統計学上の意義を否定するものではないが、実際の心理学研究において適用された場合には、非現実的で無理な前提条件を設けなければならないという大きな問題が生じてくる。

③ 変量効果モデルが考慮されてよい条件下であっても、実験計画および分散分析の本質を正しく理解しているならば、固定効果モデルによっても実質的に十分なデータ解析を行うことができる。

引用文献

遠藤健治 2002 例題からわかる心理統計学 東京：培風館

後藤宗理・大野木裕明・中澤 潤編著 2000 心理学マニュアル 要因計画法 京都：北大路書房
南風原朝和 2002 心理統計学の基礎 総合的理解のために 東京：有斐閣
肥田野直・瀬谷正敏・大川信明 1961 心理・教育統計学 東京：培風館
生沢雅夫 1977 実験計画（社会科学・行動科学のための数学入門 7） 東京：新曜社
岩原信九郎 1965 新訂版 教育と心理のための推計学 東京：日本文化科学社
岩原信九郎 1967 増補版 推計学による新教育統計法 東京：日本文化科学社
海保博之編著 1985 心理・教育データの解析法10講 基礎編 東京：福村出版
倉智佐一・山上 暁編著 1994 改訂 要説 心理統計法 京都：北大路書房
森 敏昭・吉田寿夫編著 1990 心理学のためのデータ解析テクニカルブック 京都：北大路書房
奥野忠一・芳賀敏郎 1969 実験計画法（新統計学シリーズ2） 東京：培風館
繁舩算男・柳井晴夫・森 敏昭編著 1999 Q & A で知る統計データ解析 -DOs and DON'Ts-（心理学セミナーテキストライブラリ=3） 東京：サイエンス社
篠原弘章 1984a 統計解析（行動科学のためのBASIC第1巻） 京都：ナカニシヤ出版
篠原弘章 1984b 実験計画法（行動科学のためのBASIC第2巻） 京都：ナカニシヤ出版
篠原弘章 1986 続実験計画法（行動科学のためのBASIC第3巻） 京都：ナカニシヤ出版
田中 敏 1996 実践心理データ解析 東京：新曜社
田中 敏・山際勇一郎 1989 ユーザーのための教育・心理統計と実験計画法 東京：教育出版
鷲尾泰俊 1974 実験計画法入門 東京：日本規格協会
渡辺 洋編著 2002 心理統計の技法（シリーズ・心理学の技法） 東京：福村出版
山内光哉 1998 心理・教育のための統計学 <第2版> 東京：サイエンス社

参考文献

浅井邦二監訳 1999 改訂 心理学実験計画入門 東京：学芸社（Robert L. Solso & Homer H. Johnson An introduction to experimental design in psychology, 4th ed. 1989）

- 遠藤健治 2002 SPSSによる分散分析の手順 改訂版
東京：北樹出版
- 廣津千尋 1976 分散分析（新しい応用の数学13）
東京：教育出版
- 池田 央編 1989 統計ガイドブック 東京：新曜社
- 石村貞夫 1992 分散分析のはなし 東京：東京図書
- 石村貞夫 2002 SPSSによる分散分析と多重比較の
手順 [第2版] 東京：東京図書
- 大村 平 1984 実験計画と分散分析のはなし 東
京：日科技連出版社
- Kirk, R. E. 1995 Expeimental design : Procedures for
behavioral science (3 rd ed.) Belmont :Brooks/Cole
- 小牧純爾 2000 心理学実験の理論と計画 京都：ナ
カニシヤ出版
- 芝 祐順・渡部 洋・石塚智一編 1984 統計用語辞
典 東京：新曜社
- 豊田秀樹 1994 違いを見ぬく統計学 実験計画と分
散分析入門（ブルーバックスB-1013） 東京：講談社
- Winer, B. J. 1971 Statistical principles in
experimental design (2 nd ed.) New York : McGraw-
Hill